PAPER

Web-Browsing QoE Estimation Model

Toshiko TOMINAGA†a), Kanako SATO†, Noriko YOSHIMURA†, Masataka MASUDA††, Hitoshi AOKI†, *and* Takanori HAYASHI†, *Members*

SUMMARY Web browsing services are expanding as smartphones are becoming increasingly popular worldwide. To provide customers with appropriate quality of web-browsing services, quality design and in-service quality management on the basis of quality of experience (QoE) is important. We propose a web-browsing QoE estimation model. The most important QoE factor for web-browsing is the waiting time for a web page to load. Next, the variation in the communication quality based on a mobile network should be considered. We developed a subjective quality assessment test to clarify QoE characteristics in terms of waiting time using 20 different types of web pages and constructed a web-page QoE estimation model. We then conducted a subjective quality assessment test of web-browsing to clarify the relationship between web-page QoE and web-browsing QoE for three web sites. We obtained the following two QoE characteristics. First, the main factor influencing web-browsing QoE is the average web-page QoE. Second, when web-page QoE variation occurs, a decrease in web-page QoE with a huge amplitude causes the web-browsing QoE to decrease. We used these characteristics in constructing our web-browsing QoE estimation model. The verification test results using non-training data indicate the accuracy of the model. We also show that our findings are applicable to web-browsing quality design and solving management issues on the basis of QoE.

key words: *QoE, waiting time, web page, web browsing, subjective quality assessment test, estimation model*

1. Introduction

Web browsing services are expanding as smartphones are becoming increasingly popular worldwide [1]. It has been reported that web-page load time has an impact on user behavior. Nielsen reported three important limits to web response time [2]. 0.1 seconds is about the limit at which a user feels that a system is reacting instantaneously. One second is about the limit at which a user's flow of thought will remain uninterrupted, even though the user will notice the delay. Ten seconds is about the limit for keeping a user's attention focused on dialogue. Amazon calculated that a page-load slowdown of just one second could cost it $1.6 billion in sales each year [3]. Google reported that half a second delay caused a 20% drop in traffic [4].

To provide customers with appropriate quality of web-browsing services, quality design and in-service quality management on the basis of quality of experience (QoE) [5] is important. We usually browse multiple web pages, not just a single web page. For smartphone users, web-browsing QoE is affected by mobile-network quality variation. Therefore, a web-browsing QoE estimation model must take into account both multiple web-page browsing and network quality variation.

The International Telecommunications Union (ITU-T) has published several recommendations for web-browsing QoE. The QoE factors affecting web-browsing are described in ITU-T Rec. G.1031 [6]. Web-browsing QoE is dependent on various factors that are related to users, context, and systems. The most important QoE factor of web-browsing is the waiting time for a web page to load. A subjective testing methodology for web-browsing is described in ITU-T Rec. P.1501 [7]. The recommendation provides guidance on the selection of a test environment, equipment, and content and test procedures concerning participants, tasks, and quality measures. Obtaining web-browsing QoE characteristics and constructing QoE estimation models based on these recommendations is necessary for QoE design and QoE management.

Much research on QoE characteristics based on waiting time has been conducted. Delleart et al. evaluated the QoE difference among not waiting and waiting with or without information when using Internet magazines [8]. The key result is that users become annoyed by waiting, but this negative feeling does not carry over to the evaluation of web content. However, web pages with an unknown waiting time have a negative carry-over effect on web-browsing. Galletta et al. evaluated performance measures from three viewpoints: web-site satisfaction, behavioral intentions, and number of complete tasks [9]. Participants evaluated the above three viewpoints for the web search task with a delay of between 0–12 s. The QoE decreased rapidly when the waiting time was less than 4 seconds but saturated over 8 seconds. Niida et al. developed a web assessment system in the field and conducted a QoE evaluation of waiting times for the top of a website, file download, e-mail, and so on [10]. These reports showed the importance of waiting time for web-page QoE, and that QoE characteristics depend on the type of web service.

If we carry out in-service quality management based on QoE, a real-time waiting-time measurement method and QoE estimation models are necessary. Quality-of-experience estimation models for single web pages or for an

Manuscript received October 26, 2016.
Manuscript revised February 24, 2017.
Manuscript publicized March 29, 2017.

†The authors are with the NTT Network Technology Laboratories, NTT Corporation, Musashino-shi, 180-8585 Japan.
††The author is with the NTT Information Network Laboratory Group, NTT Corporation, Musashino-shi, 180-8585 Japan.
a) E-mail: toshiko.tominaga@lab.ntt.co.jp

DOI: 10.1587/transcom.2016EBP3411

application have been proposed [9]–[14]. However, the waiting time conditions for these models have different ranges, so QoE comparison based on a unified waiting-time range cannot be conducted.

There are many types of web pages. We need to know each web-page QoE characteristic, but it is impossible to obtain them all. Therefore, classification of web-page QoE characteristics and obtaining the QoE characteristics of several types of web services classified on the basis of the same waiting-time range are necessary.

Web-browsing QoE estimation models using the search task with two web pages are discussed in ITU-T G.1030 Appendix II [15]. Hoßfeld et al. proposed a web-browsing QoE estimation model based on web-page QoE that uses a hidden memory Markov model (HMMM). This model estimates the QoE of each page intended for web-browsing of over 40 pages [16]. Web pages consist of randomly chosen image downloads. Therefore, it is not clear whether this model is also applicable to general web-site browsing.

In this paper, we propose a web-browsing QoE estimation model. We investigated several types of web-page browsing, such as general use, and set the web-browsing time to 1–2 minutes considering the typical use for a certain purpose. The web-browsing time is the time recommended by ITU-T Rec. P.1501 [7]. The waiting time for each web-page is the input data of this model. First, we obtained the web-page QoE of the waiting time for 20 types of web pages to clarify web-page QoE characteristics through a subjective quality assessment test. We found two categories of web-page QoE characteristics and constructed web-page QoE estimation models. We then clarified the relationship between web-page QoE and web-browsing QoE for three web sites through a subjective quality assessment test. From the results, we constructed our web-browsing QoE estimation model. We verified the proposed model with non-training data, and it showed good accuracy. Finally, we applied our findings to quality design and management issues for web-browsing systems and services.

The remainder of this paper is structured as follows. In Sect. 2, we clarify the web-page QoE characteristics for waiting time through a subjective quality assessment test. Then, we introduce a web-page QoE estimation model we constructed that is based on these characteristics. In Sect. 3, we clarify the relationship between web-page QoE and web-browsing QoE. In Sect. 4, we verify the proposed web-browsing QoE estimation model by using non-training data. The QoE design and management issues using our model are discussed in Sect. 5. Finally, we conclude this paper in Sect. 6.

2. Web-Page QoE Characteristics and QoE Estimation Model

In this section, we first clarify the web-page QoE characteristics for waiting time. We conducted a subjective quality assessment test for 20 types of web pages. Next, we categorized thee of web pages' QoE characteristics using cluster analysis. Finally, we constructed a web-page QoE estimation model.

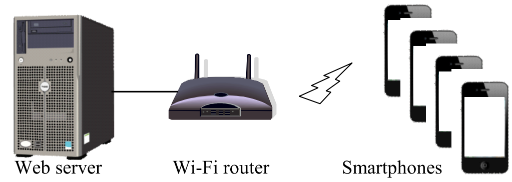

Fig. 1 Web-page QoE test environment.

Table 1 Test environment conditions for web-page QoE.

Smartphone	SH-01F OS: Android 4.2.2 Memory: 2 GB of SRAM, 32 GB of ROM CPU: MSM8974 with 2.2-GHz quad-core Display size: 5.0 inches Display resolution: 1080 × 1920 Display color: about 16.78 million
Web browser	Browser (pre-installed) Version: 4.4.2-01.00.02
Web server	BTO OS: Windows 7 Ultimate 32-bit Memory: 12 GB CPU: Core i7-960 with 3.2-GHz quad-core Apache HTTP Server: Version 2.4.9
Wi-Fi Router	WHR-300HP

Table 2 Quality assessment conditions for web-page QoE.

No. of web pages	20
Waiting time (s)	1.2, 3, 4, 6, 8.5, 12, 20
Assessment method	DCR
Participants Avg. age	24 (12 males and 12 females) 34.9
Room illuminance	about 300 lx
Display distance	Free (not fixed)

2.1 Subjective Quality Assessment Test for Web Pages

We conducted a subjective quality assessment test to clarify web-page QoE characteristics for waiting time. The test environment is shown in Fig. 1, and device conditions are listed in Table 1. Participants used a smartphone to access a web server via a Wi-Fi router.

The assessment conditions are listed in Table 2. There is a huge number of web pages in the world; thus, it is difficult to evaluate them all. Therefore, we investigated Internet usage [17], [18] and selected the 20 types of web pages that are used often, as listed in Table 3. Specifically, we referred to websites written in Japanese and created web pages for the subjective quality assessment test.

We constructed web pages from php files and inserted a "latency" parameter into the files. Regarding the web-page waiting-time range, Butkiewicz et al. [19] and Nakano et al. [20] reported that the waiting time of almost all web pages is less than 20 seconds; therefore, we set 20 seconds as the maximum waiting time for web pages. When we set "latency=0" in the php file, the waiting time to display was 1.2 seconds. For that purpose, we set the latency parameter

Table 3 Content of web page.

1	Service login	11	Photo posting
2	Bank login	12	Video posting
3	Shopping login	13	Credit card authentication
4	News	14	Transfer authentication
5	Weather forecast	15	Member registration
6	Transit information	16	Photo download
7	Route search	17	Music download
8	Product search	18	Video download
9	News search	19	Mail sending
10	Comment posting	20	Service logout

Table 4 Five-grade impairment scale.

Score	Rating scale
5	Waiting time is imperceptible
4	Waiting time is perceptible but not annoying
3	Waiting time is slightly annoying
2	Waiting time is annoying
1	Waiting time is very annoying

to values of 0, 1.8, 2.8, 4.8, 5.3, 10.8, and 18.8 seconds so that the quality assessment conditions for waiting time were 1.2, 3, 4, 6, 8.5, 12, and 20 seconds.

A degradation quality rating (DCR) method [21] was used for evaluating the waiting time quality of a web page. A five-grade impairment scale is listed in Table 4. The rating scale of DCR is adequate for rating perception of waiting time. Each participant evaluated the waiting time for test web pages compared with that of daily use.

Twenty-four participants aged 21–49 (12 males and 12 females) participated in the test, and the average age was 34.9. They were familiar with using smartphones but were non-experts who were completely unfamiliar with the technical behavior of the equipment under test. Four participants simultaneously evaluated the QoE for the 20 types of web pages, including the waiting time. Both quality assessment conditions, the webpage content and the waiting time, were presented to participants randomly. However, the same random order was presented to all participants in groups of four persons. Subjective quality assessment test was conducted in a room where the brightness was about 300 lx [22], and the distance to the display of the smart phone was taken to be the distance at which a smart phone is normally used and was not fixed.

2.2 Web-Page QoE Characteristics and QoE Classification

The mean opinion score obtained from the DCR method (DMOS) was calculated as a web-page QoE. Figure 2 shows the five major QoE characteristics of all 20 types of web pages. The QoE of music download has a higher QoE than that of weather forecast.

The ward method [23], which is a cluster analysis method, was applied for categorizing web pages based on the QoE characteristics for waiting time. Figure 3 shows the 20 types of web-page QoE dendrograms obtained with this analysis method [24]. The horizontal axis indicates the squared Euclidean distance of each cluster. The small

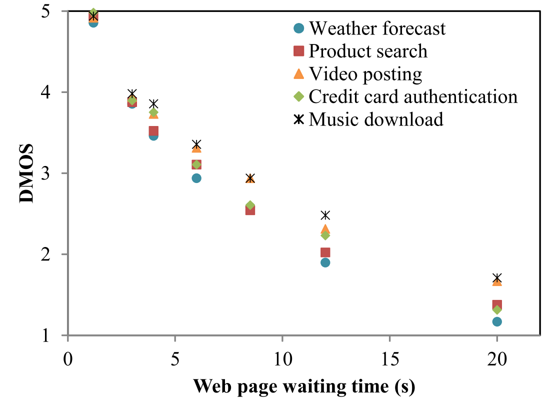

Fig. 2 Relationship between waiting time and web-page QoE (i.e., DMOS).

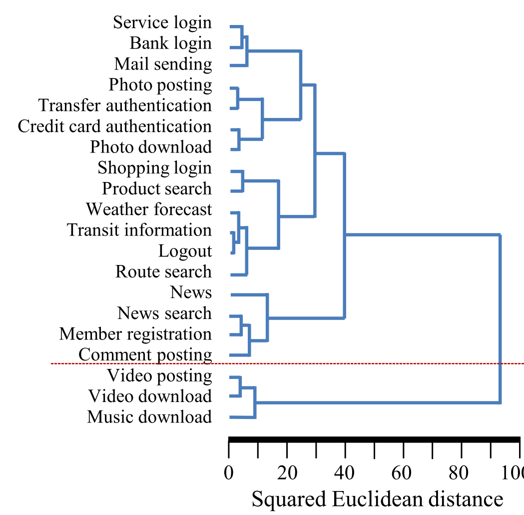

Fig. 3 Cluster analysis result.

squared distance showed high similarity between the two QoE characteristics, so they connected. The number of categories can be determined by the squared Euclidean distance. We determined the number of categories by using the squared Euclidean distance to group characteristics and checking for significant differences between the groups. As a result, we found two categories. We call one category model "large-data," which includes three web-page QoE characteristics: video posting, music downloading, and video downloading. The other category model is called "not-large-data," which includes 17 types of web pages other than the 3 mentioned above. The web-page QoE characteristics of the large-data category model showed tolerance to waiting time. The not-large-data category model characteristics showed a low QoE tendency.

2.3 Web-Page QoE Estimation Models

We constructed two web-page QoE (i.e, DMOS) estimation models for the two categories as a function of waiting time (t). The large-data category model is shown in Eq. (1), and the not-large-data category model is shown in Eq. (2).

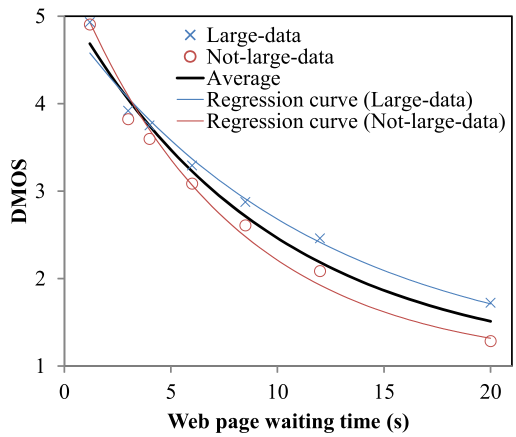

Fig. 4 Web-page QoE characteristics of two categories.

Table 5 Estimation accuracy for web-page QoE models.

Web-page QoE model	PCC	RMSE
Category models:		
a) Large-data	0.99	0.15
b) Not-large-data	0.99	0.14
Average category model	0.99	0.12

These regression lines are also described in Fig. 4. If we have information on a web page's category, we can use the two category models with good accuracy. If we cannot obtain the category information, we can use the average model of the two categories, i.e., the web-page QoE average category model. The web-page QoE average category model is shown in Fig. 4 and written as Eq. (3).

$$DMOS = 3.967exp(-t/11.628) + 1.000, \quad (1)$$

$$DMOS = 4.618exp(-t/7.463) + 1.000, \quad (2)$$

$$DMOS = 4.179exp(-t/9.524) + 1.000, \quad (3)$$

where t is waiting time and a $DMOS$ over 5 is clipped by 5.

Table 5 shows the web-page QoE estimation models' accuracies. It shows the Pearson product-moment correlation coefficient (PCC) between the measured DMOS and estimated DMOS and the root mean square error (RMSE) with each web-page QoE estimation model. The PCC was 0.99 and showed a higher correlation. We compared the RMSE with the mean of the 95% confidence interval (MCI) for subjective DMOS. The subjective DMOS score varied among the participants, even though the test condition was the same. We determined that the model had a high accuracy when the RMSE was less than the MCI. The MCI of this subjective quality assessment test was 0.16. There was a significant difference in the DMOS characteristics between the large- and not-large-data category models. However, the average category model's RMSE was also less than the MCI and exhibited high estimation accuracy. Therefore, we conclude that the category average model is suitable as a web-page QoE estimation model.

3. Web-Browsing QoE Characteristics and Proposed Web-Browsing QoE Estimation Model

We constructed our proposed web-browsing model by using web-page QoEs. First, we conducted a subjective quality assessment test and obtained the relationship between web-browsing QoE and web-page QoE. Next, we constructed our web-browsing QoE estimation model using the measured web-page QoEs. Finally, we re-learned the coefficients for web-browsing QoE estimation by using the web-page QoE average category model for each web page's waiting time.

3.1 Subjective Quality Assessment Test for Web-Browsing

We conducted a subjective quality assessment test to clarify the QoE characteristics for web-browsing waiting time. The test environment and device conditions are the same as in Fig. 1 and Table 1.

The ITU-T Rec. P.1501 recommended that typical web pages should be chosen when selecting content. Content examples are chosen to reflect typical web pages frequently browsed by typical Internet users. Three web sites, a social networking service (SNS) site, shopping site, and portal site, were selected as typical web pages [17], [18]. We made the web-pages and web scenarios listed in Table 6. The type of web page, which includes both large- and not-large-data category, was obtained from cluster analysis. We set seven waiting time conditions: 1.2, 3, 4, 6, 8.5, 12, and 20 seconds.

Web browsing time conditions must include multiple request-response patterns, that is, multiple web pages [25]. Therefore, P.1501 has recommended that the web-browsing time be set between 1–2 minutes [7]. The maximum waiting time for a web page was 20 seconds in this experiment, so the number of web pages in web-browsing was set to be equal to or less than six pages. We ignored the page holding time during web-browsing.

The subjective quality assessment test consisted of two sessions. In session I, participants evaluated the web-page QoE of web-page waiting time for the three web sites by using a 5-grade DCR method, the same as in Sect. 2. For the SNS site, participants evaluated 42 conditions, i.e., the 7 waiting time conditions for each of the 6 pages. Similarly, 42 conditions for the shopping and 35 conditions for the portal site were evaluated.

In session II, a web-browsing QoE evaluation test was conducted, and the participants evaluated task-dependent web-browsing for the three web sites. A five-point absolute category rating (ACR) method was used for web-browsing evaluation because, for the quantifying web-browsing QoE, we considered service satisfaction through web-browsing. The quality scales of ACR are shown in Table 7.

A total of 56 conditions for each site were used. That is, the 7 conditions were the same waiting time conditions for all web pages, and 49 conditions were assigned different waiting time conditions for each web page by orthogonal array. The test conditions are listed in Table 8.

Table 6 Web scenarios for training data.

SNS site

Page	Content
1	Top page
2	Video list
3	Video posting
4	Comment posting
5	Photo list
6	Photo posting

Shopping site

Page	Content
1	Top page
2	Product search
3	Product information
4	Add cart
5	Buy
6	Music download

Portal site

Page	Content
1	Top page
2	Product support
3	Model name search
4	Product manual list
5	Product manual download

Table 7 Five-grade quality scale.

Score	Rating scale
5	Excellent
4	Good
3	Fair
2	Poor
1	Bad

Table 8 Test conditions for training data.

Session I: Web-page QoE

Content	SNS	Shopping	Portal
Assessment method	DCR		
Waiting time (s)	1.2, 3, 4, 6, 8.5, 12, 20		
No. of web pages	6	6	5
No. of conditions	42	42	35

Session II: Web-browsing QoE

Content	SNS	Shopping	Portal
Assessment method	ACR		
Total waiting time (s)	7.2–120	7.2–120	6.0–100
No. of conditions	56	56	56

Common conditions

Room illuminance	about 300 lx
Display distance	about 30–40 cm (Smartphone stand)
Participants	24 (12 males and 12 females)
Avg. age	23.8

Twenty-four participants aged 20–38 (12 males and 12 females) participated in the test and the average age was 23.8. They were familiar with using smartphones to access SNS, shopping, and portal sites and were completely unfamiliar with the technical behavior of the equipment under test.

3.2 Web-Browsing QoE Characteristics and QoE Influencing Factors

The relationship between web-browsing QoE and web-page QoE is shown in Fig. 5. The Y-axis shows the MOS of web-browsing QoE for each 56 conditions of three web sites. The X-axis shows the average web-page QoEs ($DMOS_{avg}$) corresponding to the 56 conditions. $DMOS_{avg}$ is given as

$$DMOS_{avg} = \frac{1}{n}\sum_{i=1}^{n} DMOS_i, \tag{4}$$

where $DMOS_i$ means the DMOS for the waiting time for the ith page and n represents the number of web pages.

The MOS regression line and equation using $DMOS_{avg}$ are also shown in Fig. 5. The PCC between the measured MOS and $DMOS_{avg}$ was 0.89. As a result, we found that $DMOS_{avg}$ was the dominant influencing factor for estimating MOS. However, the RMSE of 0.32 was larger than the MCI of 0.29, so the estimation error with $DMOS_{avg}$ was not small enough.

We conducted an error analysis to clarify the other quality influencing factor. The $DMOS_{avg}$ was a dominant factor for estimating web-browsing QoE, but we did not consider web-page QoE variations in web-browsing QoE. Therefore, we created another quality index, $DMOS_{var}$, as a measure of web-page QoE variation. The definition of $DMOS_{var}$ is given as

$$DMOS_{var} = \frac{1}{n-1}\sum_{i=1}^{n-1} |DMOS_{i+1} - DMOS_i|. \tag{5}$$

We also used a different quality index $DMOS_{wrs}$, which means that decreasing web-page QoE decreases web-browsing QoE. An example was estimated in a similar manner for long-video QoE from the time-series data of multiple short video QoEs for a video streaming service [26]. The paper explained that the decreasing short-video QoE with a huge amplitude seriously impairs long-video QoE. In other words, for web browsing, too, the change from a state of poor quality to a state of good quality has little effect on overall service quality, and the quality index $DMOS_{wrs}$ was defined as Eq. (6), considering the effect of a change from a state of good quality to a state of poor quality.

$$DMOS_{wrs} = \frac{1}{n-1}\sum_{i=1}^{n-1} |DMOS_{i+1} - DMOS_i| \cdot S, \tag{6}$$

$$\begin{cases} S = 0, if (DMOS_{i+1} - DMOS_i) \geq 0 \\ S = 1, if (DMOS_{i+1} - DMOS_i) < 0. \end{cases}$$

Figure 6 shows the relationship between the estimation error and $DMOS_{wrs}$. The X-axis shows the estimation error with MOS estimation using $DMOS_{avg}$. The Y-axis shows the $DMOS_{wrs}$. When $DMOS_{wrs}$ is large, the measured MOS is lower than the estimated MOS. The PCC between

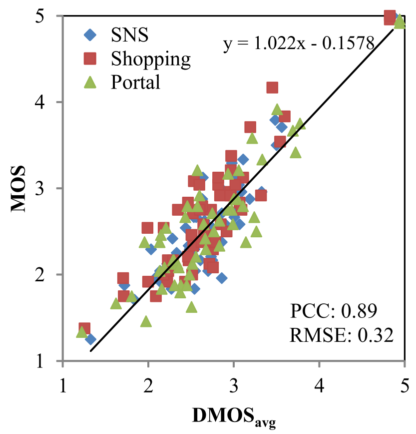

Fig. 5 Relationship between web-browsing QoE (MOS) and average web-page QoE ($DMOS_{avg}$).

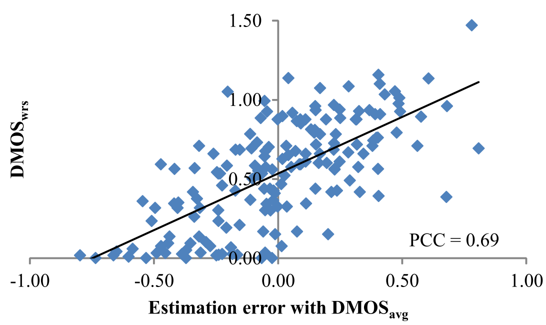

Fig. 6 Relationship between estimation error and $DMOS_{wrs}$.

the $DMOS_{wrs}$ and estimation error was 0.69. This is because a decreasing web-page QoE with a huge amplitude seriously impairs web-browsing QoE. In addition, the PCC was 0.66 when using $DMOS_{var}$.

As a result, we obtained the following two QoE characteristics. The main factor influencing web-browsing QoE is the average web-page QoE. The second is the web-page QoE variation index. We constructed and evaluated four web-browsing QoE estimation models by using these characteristics. The four models are listed in Table 9. A MOS over 5 was clipped by 5. The models included two web-page QoE variation indexes, $DMOS_{var}$ and $DMOS_{wrs}$, and the interaction between the average web-page QoE and its variation measure was considered.

The four models' estimation results are listed in Table 10. A good PCC (0.95) was achieved between the measured and estimated MOS for all models. We chose the model that had statistically significant explanatory variables. Hence, we concluded that model III, consisting of $DMOS_{avg}$ and $DMOS_{wrs}$ (proposed model), is suitable for estimating web-browsing QoE. The relationship between measured and estimated MOS is shown in Fig. 7.

3.3 Web-browsing QoE Estimation Model by Training Data

To estimate the web-browsing QoE by using each web page's

Table 9 Four web-browsing QoE estimation models.

Model	Equation
I	$a_0 + a_1 DMOS_{avg} + a_2 DMOS_{var}$
II	$a_0 + a_1 DMOS_{avg} + a_2 DMOS_{var} + a_3 DMOS_{avg} \cdot DMOS_{var}$
III	$a_0 + a_1 DMOS_{avg} + a_2 DMOS_{wrs}$
IV	$a_0 + a_1 DMOS_{avg} + a_2 DMOS_{wrs} + a_3 DMOS_{avg} \cdot DMOS_{wrs}$

Table 10 Four models' coefficients and their estimation accuracies.

Model	a_0	a_1	a_2	a_3	PCC	RMSE
I	0.051	1.119	-0.400	-	0.95	0.23
II	-0.007	1.142	-0.322	-0.029	0.95	0.23
III	-0.010	1.118	-0.735	-	0.95	0.22
IV	-0.079	1.145	-0.533	-0.077	0.95	0.22

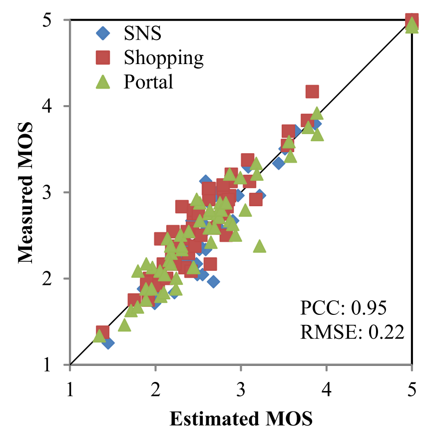

Fig. 7 Web-browsing QoE estimation using measured DMOS.

waiting time, we integrated web-page QoE average category model into model III. First, the DMOS is calculated from the web page waiting time using Eq. (3). Next, Eq. (4) and Eq. (6) are respectively used to calculate $DMOS_{avg}$ and $DMOS_{wrs}$. The result is then used to re-learning the coefficients of model III. As a result, our proposed web-browsing QoE estimation model is given by the following Eq. (7).

$$MOS = -0.766 + 1.176 DMOS_{avg} - 0.623 DMOS_{wrs}, \tag{7}$$

where a MOS over 5 is clipped by 5. The relationship between the measured MOS and estimated MOS for training data is shown in Fig. 8. As a result, the model using re-learning coefficients exhibited a higher PCC and lower RMSE than when using the measured DMOS.

4. Verification of Proposed Web-Browsing QoE Estimation Model

To verify the proposed web-browsing QoE estimation model, we used data from two subjective quality assessment tests. The test environment and device conditions were the same as shown in Fig. 1 and Table 1. The test for news and SNS sites, and their web scenarios were different from the training

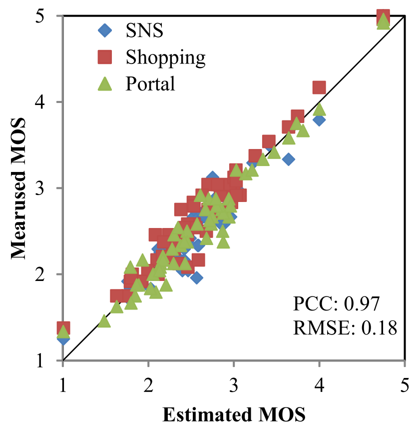

Fig. 8 Web-browsing QoE estimation using waiting time.

Table 11 Web scenarios for non-training data.

News site

Page	Content
1	Top page
2	Article 1
3	Article 2

SNS site

Page	Content
1	Top page
2	Friend search
3	Friend registration
4	Comment posting
5	Photo list
6	Photo posting

Shopping site

Page	Content
1	Top page
2	Product search
3	Product information
4	Add cart
5	Buy
6	Music download

data. The other test for the shopping site, and its web scenario was the same as the training data, but the participants were different from those of the training data.

Web scenarios are shown in Table 11. The assessment method was ACR, which was the same method as that of the training data. The ranges of the waiting time were also the same as in the training data. The test conditions are listed in Table 12.

The relationship between the MOS estimated using our web-browsing QoE estimation model and the measured MOS is shown in Fig. 9. A good PCC (0.95) was obtained between the measured MOS and estimated MOS, and the RMSE (0.24) of our model was less than the MCI (0.28) of the measured MOS. Therefore, the proposed model has a practical level of estimation accuracy. Nevertheless, differences in the estimation accuracy with content were observed and consideration of those differences is an issue for future work. We conclude that the proposed web-browsing QoE estimation model has practical and sufficient accuracy.

Table 12 Test conditions for non-training data.

Contents	News	SNS	Shopping
Assessment method	ACR		
Total waiting time (s)	3.6–60	7.2–120	7.2–120
No. of conditions	56	56	56
Participants (Avg. age)	24 (34.9)	24 (34.9)	24 (22.1)
Room illuminance	about 300 lx		

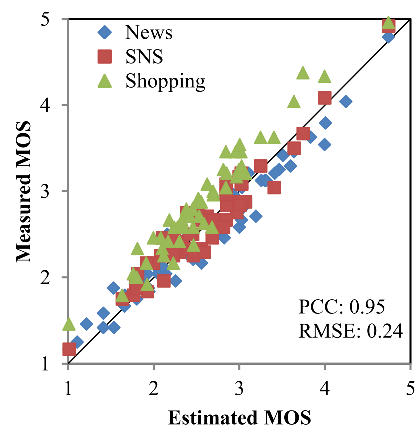

Fig. 9 Web-browsing QoE estimation using non-training data.

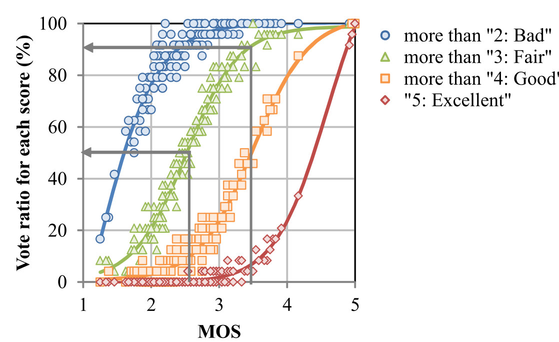

Fig. 10 Relationship between MOS and vote ratio on 5-grade quality scale.

5. Application to QoE Design and Management

In this section, we discuss the quality design/management issues for web-browsing services based on the proposed model. Figure 10 shows the measured MOS of the training data and their vote ratio to the 5-point quality scale. A MOS of 3.5 means a vote ratio greater than or equal to "3: Fair," which is 90%, and a vote ratio less than or equal to "2: Bad" is 10%. Similarly, a MOS of 2.5 means a 50% vote ratio for greater than or equal to "3: Fair." When the MOS is designed to be 3.5 for web-browsing QoE as a quality target, 90% of users choose a score of 3 or higher.

As another use case, real-time QoE measurement is useful for quality management. It is important for mobile network operators to understand the state of service quality that is being provided and to expand network facilities and change settings accordingly. A feature of mobile communications is that quality varies greatly spatially and temporally, so it is necessary to monitor for time periods in which QoE is poor in a given region or when the QoE in one region is poor relative to other regions. If the page waiting time for each web page can be measured in browsing, the method we propose can be used to visualize web browsing QoE. For example, the navigation timing API technique [27] is useful for measuring the waiting time for web pages. Therefore, web-browsing QoE is an effective quality management indicator.

6. Conclusion

We proposed a web-browsing QoE estimation model based on a web-browsing time of 1–2 minutes. We first conducted a subjective quality assessment test to clarify the QoE characteristics as determined from web-page waiting time. Twenty types of web pages were categorized from the QoE characteristics of the large- and not-large-data category models by using cluster analysis. However, the average model of the two categories had an adequately small RMSE, so we concluded that the average category model was suitable as a web-page QoE estimation model. We then conducted a subjective quality assessment test to clarify the relationship between web-page QoE and web-browsing QoE for three web sites. We developed a web-browsing QoE estimation model expressed by two parameters; the average web-page QoE and a decrease in web-page QoE as an index of web-page QoE variation. The proposed web-browsing QoE estimation model was verified with non-training data, and it showed a good accuracy. This web-browsing QoE estimation model can be useful for QoE design and management, for which we showed use cases.

Topics for further study include QoE measurement in the field and studying the relationship between web-browsing QoE and user behavior [28].

References

[1] CISCO, “Cisco visual networking index: Forecast and methodology, 2015-2020,” June 2016.

[2] J. Nielsen, Usability Engineering, Chapter 5, Morgan Kaufmann Publishers Inc., 1993.

[3] K. Eaton, “How one second could cost Amazon $1.6 billion in sales,” http://www.fastcompany.com/1825005/how-one-second-could-cost-amazon-16-billion-sales, March 2012, accessed Sept. 28, 2016.

[4] G. Linden, “Geeking with Greg,” http://glinden.blogspot.jp/2006/11/marissa-mayer-at-web-20.html, Nov. 2006, accessed Sept. 28, 2016.

[5] ITU-T Rec. P.10/G.100, “Definition of quality of experience (QoE),” Amd. 1, Jan. 2007.

[6] ITU-T Rec. G.1031, “QoE Factors in Web-browsing,” Feb. 2014.

[7] ITU-T Rec. P.1501, “Subjective testing methodology for web browsing,” Feb. 2014.

[8] B.G.C. Dellaert and B.E. Kahn, “How tolerable is delay? Consumers’ evaluations of internet websites after waiting,” Journal of Interactive Marketing, vol.13, no.1, pp.41–54, 1999.

[9] D.F. Galletta, R. Henry, S. McCoy, and P. Polak, “Web site delays: How tolerant are users?,” Journal of the Association for Information Systems, vol.5, no.1, pp.1–28, Jan. 2004.

[10] S. Niida, S. Uemura, and H. Nakamura, “Mobile services—User tolerance for waiting time—,” IEEE Veh. Technol. Mag., vol.5, no.3, pp.61–67, Sept. 2010.

[11] S. Egger, P. Reichl, T. Hoßfeld, and R. Schatz, “Time is Bandwidth? Narrowing the gap between subjective time perception and quality of experience,” IEEE ICC2012, Ottawa, Canada, June 2012.

[12] M. Fiedler, T. Hoßfeld, and P. Tran-Gia, “A generic quantitative relationship between quality of experience and quality of service,” IEEE Netw., vol.24, no.2, pp.36–41, March/April 2010.

[13] F. Liberal, A. Ferro, and J.O. Fajardo, “PQoS based model for assessing significance of providers statistically,” HET-NETs 2005, pp.18/1–18/10, July 2005.

[14] R.E. Kooij, R.D. van der Mei, and R. Yang, “TCP and web browsing performance in case of bi-directional packet loss,” Comput. Commun., vol.33, pp.S50–S57, Nov. 2010.

[15] ITU-T Rec. G.1030, “Estimating end-to-end performance in IP networks for data applications,” Feb. 2014.

[16] T. Hoßfeld, R. Schatz, S. Biedermann, A. Platzer, S. Egger, and M. Fiedler, “The memory effect and its implications on web QoE modeling,” In 23rd International Teletraffic Congress (ITC 23), San Francisco, USA, Sept. 2011.

[17] Ministry of Internal Affairs and Communications in Japan, White paper 2016, “http://www.soumu.go.jp/johotsusintokei/whitepaper/ja/h28/pdf/index.html,” July 2016, accessed Sept. 28, 2016 (in Japanese).

[18] Quantcast, “Top Sites,” http://www.quantcast.com/top-sites-1, accessed Sept. 28, 2016.

[19] M. Butkiewicz, H.V. Madhyastha, and V. Sekar, “Understanding website complexity: Measurements, metrics, and implications,” IMC’11, pp.313–328, Nov. 2011.

[20] Y. Nakano, N. Kamiyama, K. Shiomoto, G. Hasegawa, M. Murata, and H. Miyahara, “Web performance acceleration by caching rendering results,” APNOMS 2015, Busan, Aug. 2015.

[21] ITU-T Rec. P. 911, “Subjective audiovisual quality assessment methods for multimedia applications,” Dec. 1998.

[22] Ministry of Health, Labour and Welfare, “Guidelines for industrial health controls of VDT operations,” http://www.mhlw.go.jp/houdou/2002/04/h0405-4.html, accessed Feb. 7, 2017 (in Japanese).

[23] J.H. Ward, Jr., “Hierarchical grouping to optimize an objective function,” J. Am. Stat. Assoc., vol.58, no.301, pp.236–244, March 1963.

[24] S. Aoki, “Data analysis of Black Box WWW,” http://aoki2.si.gunma-u.ac.jp/BlackBox/BlackBox.html, accessed Sept. 28, 2016 (in Japanese).

[25] S. Egger, T. Hoßfeld, R. Schatz, and M. Fiedler, “Waiting times in quality of experience for web based services,” Fourth International Workshop on Quality of Multimedia Experience (QoMEX), pp.86–96, July 2012.

[26] Y. Shen, Y. Liu, H. Yang, and D. Yang, “Quality of experience study on dynamic adaptive streaming based on HTTP,” IEICE Trans. Commun., vol.E98-B, no.1, pp.62–70, Jan. 2015.

[27] W3C Rec., “Navigation timing,” https://www.w3.org/TR/navigation-timing/, Dec. 2012, accessed Sept. 28, 2016.

[28] H. Koto, N. Fukumoto, S. Niida, H. Yokota, S. Arakawa, and M. Murata, “Users’ reaction to network quality during web browsing on smartphones,” ITC 26, Karllskrona, Sept. 2014.

Toshiko Tominaga received her B.E. and M.E. degrees in electrical engineering from the University of Electro-Communications in Tokyo in 1987 and 1989. She joined NTT Laboratories in 1989 and has been engaged in research into the quality assessment of facsimile and video. She is currently working on quality assessment and user behavior analysis for web-browsing. She received the Telecommunication Advancement Foundation Award in Japan in 2008.

Kanako Sato graduated from Japan Women's Technical High School and joined NTT Laboratories in 1986. She has been engaged in multimedia service quality assessment. She is currently researching crowd testing for web-browsing QoE.

Noriko Yoshimura graduated from Tokyo Metropolitan Fifth Commercial High School and joined NTT Laboratories in 1986. She has been engaged in QoS measurement technologies for VoIP, video, and web services. She is currently researching web quality deterioration interval estimation technology.

Masataka Masuda received his B.E. and M.E. degrees in electrical engineering from the Shibaura Institute of Technology in Tokyo in 1997 and 1999 and his Ph.D. degree in engineering from the Tokyo University of Agriculture and Technology in Japan in 2011. He joined NTT laboratories in 1999 and has been engaged in IP networks and multimedia service quality evaluation. He received the Young Investigators' Award (IEICE) in Japan in 2002 and the Technical Committee on Communication Quality's Award (IEICE) in Japan in 2004.

Hitoshi Aoki received his B.E. and M.E. degrees in electrical and electronic engineering from Utsunomiya University in 1992 and 1994. He joined NTT Laboratories in 1994 and has been engaged in the quality assessment of speech and video telecommunications. Currently, he is primarily working on quality management and user behavior analysis for web-browsing. He is a member of the Acoustical Society of Japan.

Takanori Hayashi received his B.E., M.E., and Ph.D. degrees in engineering from the University of Tsukuba, Ibaraki, in 1988, 1990, and 2007. He joined NTT Laboratories in 1990 and has been engaged in the quality assessment of multimedia telecommunication and network performance measurement methods. Currently, he is the manager of the Communication Quality Group at NTT Network Technology Laboratories. He received the Telecommunication Advancement Foundation Award in Japan in 2008 and the Telecommunication Technology Committee Award in Japan in 2012.